Research on PCEP Extension for VLAN-based Traffic Forwarding in cloud network integration

Yue Wang
The Department of Cloud and Network Operation Technology Research Institue of China Telecom Corporation Limited
Beijing, China
wangy73@chinatelecom.cn

Aijun Wang
The Department of Cloud and Network Operation Technology Research Institue of China Telecom Corporation Limited
Beijing, China
wangaj3@chinatelecom.cn

Honglei Xu
The Department of Cloud and Network Operation Technology Research Institue of China Telecom Corporation Limited
Beijing, China
xuhl6@chinatelecom.cn

Wei Wang
The Department of Cloud and Network Operation Technology Research Institue of China Telecom Corporation Limited
Beijing, China
wangw36@chinatelecom.cn

Huanan Li
The Department of Cloud and Network Operation Technology Research Institue of China Telecom Corporation Limited
Beijing, China
lihn6@chinatelecom.cn

Zhengguang Cui
The Department of Cloud and Network Operation Technology Research Institue of China Telecom Corporation Limited
Beijing, China
cuizg@chinatelecom.cn

2022 7th International Conference on Communication, Image and Signal Processing (CCISP) | 978-1-6654-5959-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/CCISP55629.2022.9974205

***Abstract*—With the development of SDN technology, PCEP has become a path computing protocol widely used in the existing network. It takes the controller as the core of the overall architecture, and can be applied to calculate constrained paths in cross layer and cross domain environments in complex networks. According to the discussion on the application of PCEP in cloud-network integration scenarios, this paper proposes a VLAN-based traffic assurance mechanism based on the extension of PCEP. It can meet the requirements of key service assurance in the native IP environment and establish a connection oriented network tunnel. Under the VLAN-based architecture based on PCEP, the operator can perform closed-loop automatic control of the network and the mechanism helps to improve the intelligent scheduling of the network and the ability of real-time perception which meets the maintenance and operation requirements like flexible architecture, comprehensive opening of capabilities and global scheduling of resources.**

Keywords—PCC, PCE, PCEP, Data Packet, VLAN, Interface

I. INTRODUCTION

As the core of the current cloud computing market competition, cloud network integration has become a development strategy actively promoted by the world's leading operators, and gradually promotes the evolution of operators to IP network intelligence. Besides, With the growth of network scale, the problem of fine-grained network management and control has become particularly prominent.

As the core component of SDN architecture, the controller can make use of the northbound interface and enable the business to conveniently call the underlying network resources and capabilities, so as to achieve the unified scheduling and management of resources [1]. The southbound interface of the controller is responsible for communicating with the underlying equipment. It shields the differences of underlying physical devices through business function abstraction thus realizing the virtualization of resources and complete the centralized control of the underlying devices.

Currently, in the IP network of the operator, the layer 2 data message only contains the source and destination address. It is a hop by hop service for connectionless state, and there is no control message for the data flow path. For better traffic awareness and path optimization, the operators usually adopt the scheme of MPLS (multi-protocol label switching) network (MPLS-TE) or IPv6 network (SRv6) to guarantee the quality of key businesses [2]. With the continuous expansion of network scale, this approach will introduce too many complex protocols into the network devices and consume a lot of computing resources for routing operations.

PCEP (Path Computation Element Communication Protocol) is a centralized routing control scheme based on TCP. Through centralized deployment of controller or PCE (path computing element) , the controller calculates the optimal path for various services and complete the separation of path calculation and path establishment forwarding functions which realizes the global routing scheduling.

II. DEMAND ANALYSIS OF END-TO-END GUARANTEE SCHEME

In terms of routing control, traditional MPLS-VPN adds LDP (Label Distribution Protocol) protocol to the original IGP (Interior Gateway Protocols) to complete the label forwarding process. The LDP does not have the ability of traffic engineering and it is hard to realize the explicit path calculation through RSVP-TE (resource reservation protocol traffic engineering). Besides the huge TE link information is difficult to expand and maintain, it is impossible to collect and monitor the network quality in real time and complete the automatic

optimization of the optimal path, so the efficiency of information interaction is low [3].

SRv6, as an newly developed technology, make full use of the mechanism of IPv6 extension header to forward traffic through the IPv6 address identification segment in the SRH (Segment Routing Header). However, due to the SRH header, there are problems such as message overhead and low utilization of network link bandwidth. In addition, SRv6 packet processing has higher requirements on the chip. It is difficult for the existing network device to support the replication and operation of the SRH with 128 bit SID (Segment ID) extension header. The smooth upgrade and evolution to SRv6 cannot be achieved, which brings great cost pressure to operators to deploy SRv6.

RFC8735 introduces traffic engineering in multi-domain and the QoS assurance based on hybrid cloud communication scenario in single domain [4][5]. Based on those scenarios, the following criteria should be met for the architecture of traffic engineering in a native IP network:

- Suitable for both inter and intra domain scenarios.
- The solution should be suitable for both native IPv4 and IPv6 traffic.
- For traffic that needs the service assurance (priority traffic), end to end traffic assurance is achieved through deterministic QoS behavior.
- The optimal path can be dynamically adjusted according to changes in network status. There is no need to reserve resources for physical links in advance.
- The control plane of the network is based on a distributed architecture and is under the centralized control.

III. PCEP EXTENSION FOR VLAN-BASED TRAFFIC FORWARDING MECHANISM

With the large scale deployment of PCEP architecture and Ethernet interfaces in native network, an end-to-end dedicated path can be constructed by using the VLAN information in the Ethernet header to guide the packet forwarding. In this context, we proposed a PCEP Extension for VLAN-based traffic forwarding mechanism (ePCEP for VTF) by adding the VLAN information contained in the Ethernet frame which can effectively utilize VLAN information in Layer 2 Ethernet frame structure and meet the traffic engineering criteria mentioned above.

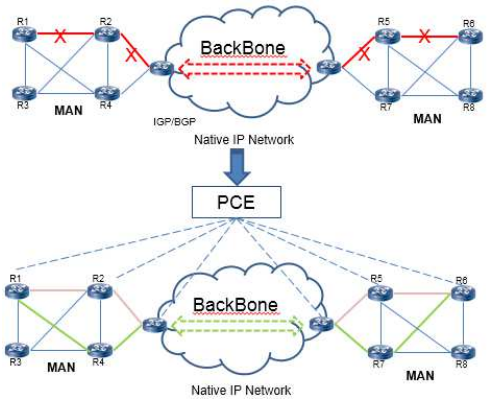

Fig. 1. The topology of intra&inter-domain scenario in native IP

The ePCEP for VTF mechanism is mainly deployed in Native IP environment and suitable for both IPv4 and IPv6. In such cases, the prefix among the underlying devices(PCCs) is distributed by BGP, no MPLS is involved.

The whole process of the traffic forwarding mechanism is based on the PCEP architecture as shown in figure 1. In the actual business scenarios, in order to distinguish different types of services, the ingress and egress PCC can deploy multiple BGP sessions [6]. Different BGP sessions distribute different prefixes and have different BGP next hops.

In the control plane, the ingress PCC can learn the different path prefixes of the source&destination peer based on the same BGP session through BGP. The controller as a PCE calculates the explicit route based on the business requirements and the traffic engineering policy and then the routing information (including source BGP peer, destination BGP peer, VLAN ID) is sent to the PCCs in the forms of PCInitiate messages. The ingress PCC responds with a PCRpt messages and forms a VLAN-Forwarding routing (VFR) table. Similarly, the VLAN-Crossing routing (VCR) table newly defined in this article will be formed by the transit&egress PCC.

In the forwarding plane, when a data packet is delivered to the ingress PCC, its source&destination address will be matched with the source&destination BGP prefix in the VFR table. If they are consistent, the data packet will be marked with the corresponding Vlan-ID label. Through the identification of the VLAN tag, the tagged packet will be sent to the PCC's specific sub-interface and further be forwarded. To the transit PCC, the packet that needs to be secured will be forwarded with a new VLAN tag relabelled. To the egress PCC, it removes the Layer 2 header that encapsulates the Vlan-ID according to the mapping information, and the packet will be forwarded in Layer 3.

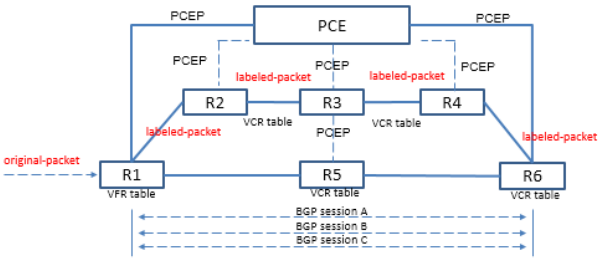

Fig. 2. Process of VLAN-based Traffic Forwarding

Figure 2 shows the process of VLAN forwarding, Generally, the main procedure of this VLAN-based traffic forwarding mechanism based on PCEP can be summarized as the following steps:

Step1: The PCE completes the end-to-end optimal path calculation of key services according to the network constraint parameters and sends the route message in the forms of PCInitiate messages to the PCCs.

Step2: The VFR table and the VCR table are generated on the ingress PCC, the transit and egress PCC respectively.

Step3: By matching with the VFR and VCR table,the packet of key services will be tagged with corresponding VLAN.

Step4: The packet tagged with VLAN will be transferred to the the PCC's specific sub-interface and then be forwarded through the VLAN tunnel.

Based on the PCEP protocol, the end-to-end traffic assurance in connection-oriented network can be achieved via the VLAN based traffic forwarding, thus the quality of communication services can be guaranteed in native IP environment. By combining it with PCE and PCC which are the elements of PCEP rather than replacing it completely, the calculation and forwarding process of the optimal route can be greatly simplified. The central controller calculates and deploys the optimal VLAN switching path (VSP) to bypass the blocked links and nodes, thus avoiding pre-reservation of resources on each network devices and achieving the overall QoS guarantee effect.

SRv6 and MPLS, as two of traffic assurance technologies widely used in the complex network at present, have the drawbacks like supports only in IPv6 environment, high cost of the protocol resource and has the obvious SRH overhead problems. Compared with them, the ePCEP for VTF mechanism, as a new traffic assurance technology, has the following advantages:

- Avoids SRH Overhead problem.
- Uses a completely new address space to bypass the already used MPLS label space, which avoids the possibility of conflict with other existing protocols and eliminates the need to consider the label overlap of the already used MPLS services in the MPLS-Native IP Mixed environment.
- Can fully utilize the existing PCEP architecture and can be deployed on both ipv4 and ipv6 networks.

IV. VLAN-BASED TRAFFIC FORWARDING PROCEDURES

For the complex network topology in the existing network, there are often multiple BGP Sessions between network devices, and different Sessions are used to carry different data service flows to achieve end-to-end traffic isolation and different levels of SLA (Service- Level Agreement) level guarantee.

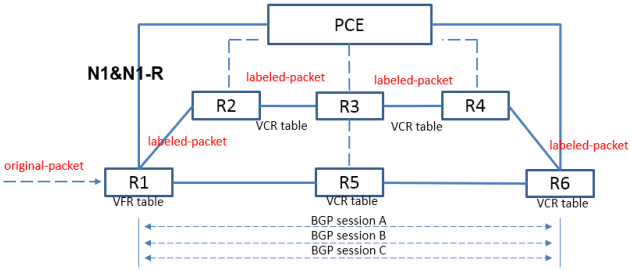

Fig. 3. VLAN-Based forwarding message announcement of ingress PCC

Figure 3 is a basic network topology under PCEP architecture. Based on the previously introduced VLAN-based traffic forwarding mechanism, the implementation process of the scheme under PCEP architecture is as follows.

As shown in the figure, three BGP instances between R1 and R6 are created through different sub-interfaces, corresponding to three BGP sessions respectively. Different BGP sessions distribute different prefixes and have different BGP next hops. In BGP session A, the source and destination IP prefix of service traffic are P1 and P6 respectively. R1 as ingress PCC can learn the path prefixes P1 and P6 of the source / destination peer based on BGP session A through BGP.

The PCE as controller and R1 as ingress PCC are connected through PCEP. Based on the global network topology and service needs, the PCE calculates the global optimization path and forms a VLAN forwarding messages which is represented as N1&N1-R in figure 3. The VLAN forwarding messages will be delivered to R1 through the message pair of PCInitiate&PCRpt. The message peers and message key parameters are defined in table 1. The R1A and R6A of SrcPeer_IP and DstPeer_IP indicate the address of R1 and R6 of BGP session A. The INTF1 indicates one of the sub-interface addresses that carries the specific service traffic which needs to be guaranteed. VLAN_R1R2 indicates the VLAN corresponding to the path from R1 to R2.

TABLE I. VLAN FORWARDING MESSAGE INFORMATION

ID	Peers	Message Type	Key Parameters
N1&N1-R	PCE-R1	PCInitiate PCRpt	Interface_Address = INTF1, SrcPeer_IP = R1A, DstPeer_IP = R6A, VLAN_ID = VLAN_R1R2

After R1 receives the VLAN forwarding messages through the PCInitiate message, a VLAN-Forwarding routing table defined below will be formed and according to the pre–learned source&destination BGP prefix and VLAN ID contained in the VLAN forwarding messages, a specific VLAN will be set up on its sub-interface. When a packet is delivered to R1, by looking up the VFR table via the source&destination IP, the corresponding VLAN tag such as VLAN_R1R2 will be labelled on the packet that needs to be secured. After that, The labelled packet will be further transferred to the sub-interface specified by VLAN.

TABLE II. VLAN FORWARDING ROUTING TABLE

Src IP	Dst IP	Sub-interface	VLAN_ID
Prefixes of R1 Session1	Prefixes of R6 Session1	INTF 1	VLAN_R1R2
Prefixes of R1 SessionX …	Prefixes of R6 SessionX …	INTF X …	VLAN X …

Based on the BGP prefix contained in the packet, the VFR table maintained in the ingress PCC, as shown in table 2, is used to identify the packet that needs to be secured.

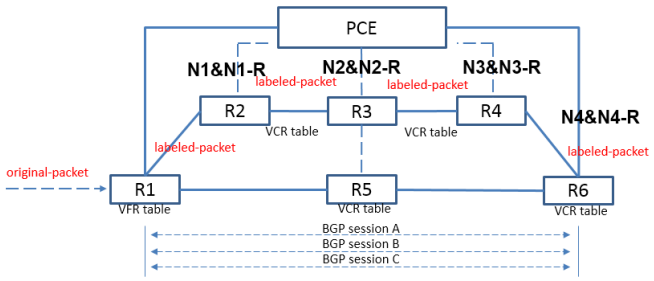

Fig. 4. VLAN-Based crossing message announcement of transit&egress PCC

Similarly, VLAN crossing messages are also formed by the PCE when the PCE calculates the global optimization route. The transit PCC like R2, R3, R4 and egress PCC like R6 will receive the VLAN crossing messages which is Represented as N1&N1-R to N4&N4-R in figure 4. The VLAN crossing messages are also carried by the PCInitiate and PCRpt message pair. The message peers and message key parameters are defined in table 3. The parameters are divided into inbound and outbound. The VLAN IDs of inbound and outbound form a key-value pair which indicates a new VSP. The interface addresses indicate the inbound and out bound sub-interface addresses that carries the specific service traffic which needs to be guaranteed.

TABLE III. VLAN CROSSING MESSAGE INFORMATION

ID	Peers	Message Type	Key Parameters
N1&N1-R	PCE-R2	PCInitiate PCRpt	inbound Interface_Address = INTF1, VLAN_ID = VLAN_R1R2 outbound Interface_Address = INTF2, VLAN_ID = VLAN_R2R3
N2&N2-R	PCE-R3	PCInitiate PCRpt	inbound Interface_Address = INTF1, VLAN_ID = VLAN_R2R3 outbound Interface_Address = INTF2, VLAN_ID = VLAN_R3R4
N3&N3-R	PCE-R4	PCInitiate PCRpt	inbound Interface_Address = INTF1, VLAN_ID = VLAN_R3R4 outbound Interface_Address = INTF2, VLAN_ID = VLAN_R4R6
N4&N4-R	PCE-R6	PCInitiate PCRpt	inbound Interface_Address = INTF1, VLAN_ID = VLAN_R4R6 outbound Interface_Address = INTF2, VLAN_ID=0

After the process of VLAN-Based forwarding message announcement introduced previously, a VLAN Crossing routing table will be formed in the R2,R3,R4 as transit PCC and R6 as egress PCC. Based on the VLAN ID contained in the VLAN crossing messages, the specific VLAN will be set up on their sub-interface. When a data packet tagged with VLAN_R1R2 which is done by R1 is delivered to R2, it will look up the VCR table via tagged VLAN. If the VLAN is consistent, the ingress-VLAN as VLAN_R1R2 will be replaced with a egress-VLAN as VLAN_R2R3 by the current transit PCC. The packet labelled with new VLAN will be further delivered to the next hop.

R6, as the egress PCC, its egress-VLAN in the VCR table should be 0 which indicates it's the final hop in the whole transit process. Therefore, the egress PCC will strip the ingress-VLAN and the packet will be transited directly.

TABLE IV. VLAN CROSSING ROUTING TABLE

Ingress-Subinterface	Ingress-VLAN	Egress-Subinterface	Egress-VLAN
INTF1	VLAN_R1R2	INTF2	VLAN_R2R3
INTF3	VLAN A	INTF4	VLAN B
INTF5	VLAN C	INTF6	0

Table 4 is the VCR table maintained in transit&egress PCC. By mapping with ingress&egress VLAN, the packet to be assured will be transmitted to a predefined sub-interface and tagged with egress VLAN by transit PCC or directly forwarded by egress PCC.

So far, for all the data packets from R1 to R6, only data services matching P1 and P6 prefixes can use end-to-end guaranteed logical channels. Moreover, the path of the logical channel can be planned by the SDN controller through PCEP,

and the data entered by other interfaces will be forwarded according to the traditional routing table.

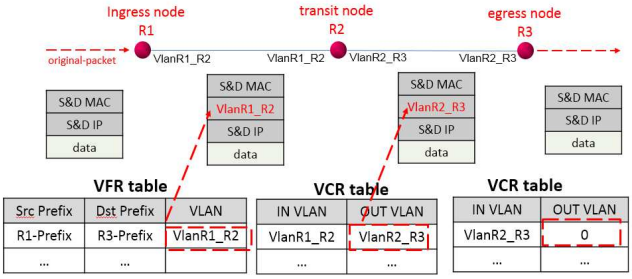

Fig. 5. Data Packet Encapsulation Process

Figure 5 shows the data packet encapsulation process of an end-to-end traffic for key application. According to IEEE 802.1Q protocol which defines the standard of data link layer, VLAN tag is usually located in Ethernet header and IP header. So in the whole process of the traffic forwarding mechanism, the VLAN tag is encapsulated between the MAC and the IP address. Based on the VFR(VLAN forward routing) table and VCR(VLAN crossing routing) table, the original packet will be transmitted along the path of the VSP through the exchange of VLAN labels.

The ePCEP for VTF mechanism meets the demands of end-to-end service assurance in native IP environment. Via calculating and deploying the optimal VSP by the central controller, the overall QoS assurance effect is achieved, and there is no more need to reserve resources for physical links in advance.

V. CONCLUSION

The large-scale deployment of Ethernet interface makes it possible to simplify the end-to-end data forwarding process by using the information contained in the layer 2 frame structure. Under this background, we proposed an ePCEP for VTF mechanism in this paper. The whole mechanism can provide end-to-end service assurance for specific customers and applications, and realize deterministic transmission of key services in IP scenarios. It makes full use of the VLAN information in layer 2 and implements full-scenario traffic access based on PCEP. The mechanism simplifies the end-to-end path calculation and forwarding process of messages while preserving the PCEP structure as much as possible. It can meet the path forwarding requirements of multi-service traffic and ensure the priority and service quality of key businesses, So as to realize flexible networking and multi-dimensional SLA path planning.

ACKNOWLEDGMENT *(Heading 5)*

Thanks to the experts from Huawei, ZTE and China Mobile for their support and constructive comments.

REFERENCES

[1] F. Paolucci, F. Cugini, A. Giorgetti, N. Sambo and P. Castoldi, "A Survey on the Path Computation Element (PCE) Architecture," in IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 1819-1841, Fourth Quarter 2013, doi: 10.1109/SURV.2013.011413.00087.

[2] I. Šeremet and S. Čaušević, "Advancing Multiprotocol Label Switching Traffic Engineering with Segment Routing in Software Defined Network environment," 2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH), 2020, pp. 1-6, doi: 10.1109/INFOTEH48170.2020.9066291.

[3] X. Chen, Y. Zhong and A. Jukan, "Multipath routing in Path Computation Element (PCE): Protocol extensions and implementation," Proceedings of the 2013 18th European Conference on Network and Optical Communications & 2013 8th Conference on Optical Cabling and Infrastructure (NOC-OC&I), 2013, pp. 75-82, doi: 10.1109/NOC-OCI.2013.6582871.

[4] Aijun Wang, Xiaohong Huang, Caixia Qou, Zhenqiang Li, Penghui Mi, “Scenarios and Simulation Results of PCE in a Native IP Network”, IETF RFC 8735, Feb 2020

[5] Aijun Wang, Boris Khasanov, Quintin Zhao, Huaimo Chen, “PCE-Based Traffic Engineering (TE) in Native IP Networks”, IETF RFC 8821, Apr 2021

[6] F. Cugini, F. Paolucci, L. Valcarenghi, P. Castoldi and A. Welin, "PCE communication protocol for resource advertisement in multi-domain BGP-based networks," 2009 Conference on Optical Fiber Communication, 2009, pp. 1-3, doi: 10.1364/OFC.2009.OWL3.